

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: MURPHY, Dennis  
REID, John

(ii) TITLE OF INVENTION: Alpha Glycosidase

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
CECCHI, STEWART & OLSTEIN  
(B) STREET: 6 BECKER FARM ROAD  
(C) CITY: ROSELAND  
(D) STATE: NEW JERSEY  
(E) COUNTRY: USA  
(F) ZIP: 07068

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5 INCH DISKETTE  
(B) COMPUTER: IBM PS/2  
(C) OPERATING SYSTEM: MS-DOS  
(D) SOFTWARE: ASCII

(vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: Unassigned  
(B) FILING DATE: Concurrently  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: HERRON, CHARLES J.  
(B) REGISTRATION NUMBER: 28,019  
(C) REFERENCE/DOCKET NUMBER: 331400-40

(ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 201-894-1700  
(B) TELEFAX: 201-994-1744

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 52 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:  
CCGAGAATTTC ATTAAAGAGG AGAAATTAAAC TATGAGAGCG CTCGTCTTTC AC

(2) INFORMATION FOR SEQ ID NO:2:

52

(i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 31 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGGAAGATCT AGGTTCCCCA TTTTCACCCC T

31

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS  
(A) LENGTH: 1,095 NUCLEOTIDES  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTG AGA GCG CTC GTC TTT CAC GGC AAC CTC CAG TAT GCC GAA ATC CCA  
Leu Arg Ala Leu Val Phe His Gly Asn Leu Gln Tyr Ala Glu Ile Pro  
5 10 15

48

AAG AGC GAA ATC CCA AAG GTC ATA GAG AAG GCA TAC ATC CCA GTC ATC  
Lys Ser Glu Ile Pro Lys Val Ile Glu Lys Ala Tyr Ile Pro Val Ile  
20 25 30

96

GAG ACA CTG ATT AAA GAA GAA ATT CCT TTT GGG CTC AAC ATA ACG GGC  
Glu Thr Leu Ile Lys Glu Glu Ile Pro Phe Gly Leu Asn Ile Thr Gly  
35 40 45

144

TAT ACC TTA AAG TTC CTC CCG AAG GAT ATT ATA GAC CTC GTT AAA GGG  
Tyr Thr Leu Lys Phe Leu Pro Lys Asp Ile Ile Asp Leu Val Lys Gly  
50 55 60

192

GGC ATC GCG AGT GAC CTG ATA GAG ATA ATC GGA ACG AGC TAC ACG CAC  
Gly Ile Ala Ser Asp Leu Ile Glu Ile Gly Thr Ser Tyr Thr His  
65 70 75 80

240

GCA ATA CTC CCC CTC CTC CCG CTT AGC AGA GTA GAA GCA CAA GTT CAG  
Ala Ile Leu Pro Leu Leu Pro Leu Ser Arg Val Glu Ala Gln Val Gln  
85 90 95

288

AGA GAT AGG GAA GTT AAG GAA GAG CTC TTC GAG CTT TCT CCA AAG GGA  
Arg Asp Arg Glu Val Lys Glu Glu Leu Phe Glu Val Ser Pro Lys Gly  
100 105 110

336

TTC TGG CTG CCA GAG CTC GCC TAT GAC CCG ATA ATC CCT GCC ATA CTG  
 phe Trp Leu Pro Glu Leu Ala Tyr Asp Pro Ile Ile Pro Ala Ile Leu  
 115 120 125

384

AAG GAC AAC GGT TAT GAG TAT CTA TTC GCC GAC GGG GAG GCG ATG CTT  
 Lys Asp Asn Gly Tyr Glu Tyr Leu Phe Ala Asp Gly Glu Ala Met Leu  
 130 135 140  
 130

432

480

145 CCA CAC CTT ATA AAG GCC CAA AGG GAA AAG CGC TTT AGG TAC ATC AGC  
Pro His 3Leu Ile Lys Ala Gln Arg Glu Lys Arg Phe Arg Tyr Ile Ser  
165 170 175

528

TAT CTC CTT GGT CTC AGG GAG CTT AGG AAG GCG ATA AAG CTC GTT TTT  
 Tyr Leu Leu Gly Leu Arg Glu Leu Arg Lys Ala Ile Lys Leu Val Phe  
 180 185 190

576

GAA GGT AAG GTA ACG CTA AAG GCA GTC AAA GAC ATC GAA GCC GTA CCC  
Glu Gly Lys Val Thr Leu Lys Ala Val Lys Asp Ile Glu Ala Val Pro  
195 200 205

624

151  
 GTT TGG GTG GCC GTG AAC ACG GCT GTA ATG CTC GGC ATC GGA AGG CTT  
 Val Trp Val Ala Val Asn Thr Ala Val Met Leu Gly Ile Gly Arg Leu  
 210 215 220  
 219

672

210 CCT CTT ATG AAT CCT AAG AAA GTG GCG AGC TGG ATA GAG GAC AAG GAC  
 Pro Leu Met Asn Pro Lys Lys Val Ala Ser Trp Ile Glu Asp Lys Asp 240  
 225 230 235

720

225 AAC ATT CTT CTA TAC GGC ACC GAT ATA GAG TTC ATT GGC TAT AGG GAC  
 Asn Ile Leu Leu Tyr Gly Thr Asp Ile Glu Phe Ile Gly Tyr Arg Asp  
 245 250 255

768

270 ATT GCA GGC TAC AGA ATG AGT GTT GAG GGA TTA TTA GAG GTT ATA GAC  
 Ile Ala Gly Tyr Arg Met Ser Val Glu Gly Leu Leu Glu Val Ile Asp  
 265 270  
 260

816

280  
GAG CTC AAC TCG GAA CTG TGC CTT CCC TCA GAG CTG AAG CAC AGT GGA  
Glu Leu Asn Ser Glu Leu Cys Leu Pro Ser Glu Leu Lys His Ser Gly  
285  
275

864

275  
 AGG GAG CTC TAC TTA CGG ACT TCG AGT TGG GCA CCA GAT AAG AGC TTG  
 Arg Glu Leu Tyr Leu Arg Thr Ser Ser Trp Ala Pro Asp Lys Ser Leu  
 295 300

912

960

AGG ATA TGG AGA GAG GAA GGG AAC GCA AGA CTT AAT ATC CTG TCC  
Arg Ile Trp Arg Glu Glu Gly Asn Ala Arg Leu Asn Met Leu Ser  
305 315 320

1008

TAC AAT ATG AGG GGC GAA CTC GCC CTT TTA GCC GAG AAC AGC GAT GCA  
Gly Glu Leu Ala Phe Leu Ala Glu Asn Ser Asp Ala  
325 330 335

1047

AGG GGA TGG GAG CCC CTC CCT GAG AGG CTG GAT GCC TTC CGG GCG  
Arg Gly Trp Glu Pro Leu Pro Glu Arg Arg Leu Asp Ala Phe Arg Ala  
340 345 350

1086

ATA TAT AAC GAT TGG AGG GGT GAA AAT GGG GAA CCT TAG  
Ile Tyr Asn Asp Trp Arg Gly Glu Asn Gly Glu Pro End  
355 360 365

*Part B1*

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 364 AMINO ACIDS
- (B) TYPE: POLYPEPTIDE
- (C) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Leu Arg Ala Leu Val Phe Mis Gly Asn Leu Gln Tyr Ala Glu Ile  
5 10 15

phe Lys Ser Glu Ile Pro Lys Val Ile Glu Lys Ala Tyr Ile Pro  
20 25 30

Val Ile Glu Thr Leu Ile Lys Glu Gln Ile Pro Phe Gly Leu Asn  
35 40 45

Ile Thr Gly Tyr Thr Leu Lys Phe Leu Pro Lys Asn Ile Ile Asp  
50 55 60

Leu Val Lys Gly Gly Ile Ala Ser Asp Leu Ile Glu Ile Ile Gly  
65 70 75

Thr Ser Tyr Tyr His Ala Ile Leu Pro Leu Leu Pro Leu Ser Arg  
80 85 90

Val Glu Ala Glu Val Gln Arg Asp Arg Glu Val Lys Glu Leu  
95 100 105

phe Glu Val Ser Pro Lys Gly Phe Trp Leu Pro Glu Leu Ala Tyr  
110 115 120

Asn Pro Ile Ile Phe Ala Ile Leu Lys Asp Asn Gly Tyr Glu Tyr  
125 130 135

Leu Phe Ala Asp Gly Glu Ala Met Leu Phe Ser Ala His Leu Asn  
140 145 150

Ser Ala Ile Lys Pro Ile Lys Pro Leu Tyr Pro His Leu Ile Lys  
 155 160 165  
 Ala Gln Arg Glu Lys Arg Phe Arg Tyr Ile Ser Tyr Leu Leu GLY  
 170 175 180  
 Leu Arg Glu Leu Arg Lys Ala Ile Lys Leu Val Phe Glu Gly Lys  
 185 190 195  
 Val Thr Leu Lys Ala Val Lys Asp Ile Glu Ala Val Pro Val Trp  
 200 205 210  
 Val Ala Val Asn Thr Ala Val Met Leu Gly Ile Gly Arg Leu Pro  
 215 220 225  
 Leu Met As? Pro Lys Lys Val Ala Ser Trp Ile Glu Asp Lys Asp  
 230 235 240  
 Asn Ile Leu Leu Tyr Gly Thr Asp Ile Glu Phe Ile Gly Tyr Arg  
 245 250 255  
 Asp Ile Ala Gly Tyr Arg Met Ser Val Glu Gly Leu Leu Glu Val  
 260 265 270  
 Ile Asp Glu Leu Asn Ser Glu Leu Cys Leu Pro Ser Glu Leu Lys  
 275 280 285  
 His Ser Gly Arg Glu Leu Tyr Leu Arg Thr Ser Ser Trp Ala Pro  
 290 295 300  
 Asp Lys Ser Leu Arg Ile Trp Arg Glu Asp Glu Gly Asn Ala Arg  
 305 310 315  
 Leu Asn Met Leu Ser Tyr Asn Met Arg Gly Glu Leu Ala Phe Leu  
 320 325 330  
 Ala Glu Asn Ser Asp Ala Arg Gly Trp Glu Pro Leu Pro Gln Arg  
 335 340 345  
 Arg Leu Asp Ala Phe Arg Ala Ile Tyr Asn Asp Trp Arg Gly Glu  
 350 355 360  
 Asn Gly Glu Pro